

Nonlocal exchange correlation in screened-exchange density functional methods

Byounggak Lee* and Lin-Wang Wang

*Computational Research Division, Lawrence Berkeley
National Laboratory, Berkeley, California 94720, USA*

Catalin D. Spataru

*Center for Integrated Science and Engineering and Center
for Electron Transport in Molecular Nanostructures,
Columbia University, New York 10027, USA*

Steven G. Louie

*Department of Physics, University of California at Berkeley,
Berkeley, California 94720, USA and
Materials Sciences Division, Lawrence Berkeley
National Laboratory, Berkeley, California 94720, USA*

Abstract

We present a systematic study on the exchange-correlation effects in screened-exchange local density functional method. To investigate the effects of the screened-exchange potential in the band gap correction, we have compared the exchange-correlation potential term in the sX-LDA formalism with the self-energy term in the GW approximation. It is found that the band gap correction of the sX-LDA method primarily comes from the downshift of valence band states, resulting from the enhancement of bonding and the increase of ionization energy. The band gap correction in the GW method, on the contrary, comes in large part from the increase of the conduction band energies. We also studied the effects of the screened-exchange potential in the total energy by investigating the exchange-correlation hole in comparison with quantum Monte Carlo calculations. When the Thomas-Fermi screening is used, the sX-LDA method overestimates (underestimates) the exchange-correlation hole in short (long) range. From the exchange-correlation energy analysis we found that the LDA method yields better absolute total energy than sX-LDA method.

PACS numbers: 71.15.Mb, 71.15.Qe, 71.20.Nr

*Electronic address: bhlee@lbl.gov

I. INTRODUCTION

Searching for better exchange-correlation functionals beyond the current local density approximation[1] (LDA) or generalized gradient approximation[2] (GGA) is one of the main research topics in density functional theory[3] (DFT) development. One particular goal in this development is to overcome the band gap problem in LDA or GGA.[4, 5] Although it is often commented that the DFT is a ground state theory and that, as a result, it intrinsically cannot get the excited state band gap, strictly speaking, the band gap of semiconductors is accessible by ground state total energy calculations. This is because the quasi-particle energy of the conduction band minimum (CBM) equals to $E(N+1) - E(N)$, where $E(M)$ is the ground state energy of the system with M electrons and N is the number of the electrons in the neutral system. Similarly, the valence band maximum (VBM) energy equals $E(N) - E(N-1)$. Furthermore, if there is no singularity in the explicit exchange-correlation functional when M crosses N (e.g., LDA, GGA, and the screened-exchange functional to be discussed below), the $E(N+1) - E(N)$ [$E(N) - E(N-1)$] will be equal to the $(N+1)$ th [N th] eigenenergy in the generalized Kohn-Sham (KS) equation for an infinite bulk system. Therefore, it is quite legitimate to require a next generation DFT to give not only the accurate ground state total energy but also accurate energy band gaps. After all, the DFT single particle energy is just the total energy derivative with respect to the occupation number of the state.[6]

To go beyond the LDA and GGA, it is now widely believed that some kind of nonlocal expression in the exchange-correlation functional is necessary.[4, 5, 7] If this expression involves the wavefunctions as well as the charge density, the scheme goes beyond the current KS scheme of the DFT. Not only there is a nonlocal integral in the total energy expression, but also there is a nonlocal potential in the generalized KS (GKS) equation (on the contrary, in strict KS schemes, e.g., the exact exchange functional, there are only local potentials in the KS equation). A simple and natural nonlocal expression in the exchange-correlation functional is the screened-exchange LDA (sX-LDA) functional.[8] By screening the Hartree-Fock (HF) exchange integral, the sX-LDA method attempts to introduce part of correlation effects in the exchange hole. On the other hand, by reproducing the LDA results in homogeneous electron gas, the sX-LDA method retains many of the LDA successes for the total energy calculations. It has been shown that the sX-LDA provides very good band gaps for

common IV-IV, III-V, II-VI semiconductors and that its total energies and the resulting atomic structures of materials are similar to, and sometimes better than, those of LDA results.[8]

In order to further improve the sX-LDA method or to come up with alternative methods, it is essential to understand the successes and failures of the current sX-LDA method by comparing with more accurate methods. For the single particle eigenvalues and the related GKS equations, we compare the sX-LDA method with the GW method[9–15], especially for the nonlocal potentials. For the total energy calculations and exchange-correlation holes, we compare sX-LDA with quantum Monte Carlo simulations. Our emphasis is in the nonlocality of the potential and exchange-correlation hole, including their amplitudes, spatial ranges, and the variations at different points. We choose bulk Si as our main test system to compare with the existing quantum Monte Carlo (QMC) study.[16] We also study the origin of sX-LDA band gap corrections and the difference in absolute total energy (not just the relative total energy as used in the lattice constant and bulk modulus calculations) in comparison with LDA.

The remainder of the paper is organized as follows. In Sec. II, we review the sX-LDA method and demonstrate the improvement of the band gap. In Sec. III, we analyze the effects of screened-exchange potential on the single particle eigenvalues and compare the behavior of eigenvalues with the behavior of quasiparticle eigenvalues in GW method. The screened-exchange potential effects on the total energy is studied using the exchange-correlation hole in Sec. IV.

II. THE SCREENED-EXCHANGE METHOD

In this section we briefly review the sX-LDA method for the discussions in the following sections. More detailed description with rigorous derivation can be found in Ref[8]. The sX-LDA method is based on the generalized Kohn-Sham (GKS) schemes. Unlike in the usual Kohn-Sham (KS) schemes, where the electron-electron interaction is included only via electron density, $\rho(\mathbf{r})$, the GKS scheme allows this interaction energy to have explicit orbital dependence:

$$E_{\text{tot}} = T + E_{\text{H}}[\rho] + E_{\text{sX}}[\{\psi\}] + R[\rho] + E_{\text{ext}}[v], \quad (1)$$

where T , $E_H[\rho]$, and $E_{\text{ext}}[v]$ are the kinetic energy, direct Hartree Coulomb energy, and external potential energy, respectively, and v is an external potential. The orbital dependent electron-electron interaction is included in the screened-exchange energy term $E_{\text{sX}}[\{\psi\}]$. $R[\rho]$ is the difference between the true exchange-correlation energy and the screened-exchange energy. The present sX-LDA method uses the exchange energy with a Thomas-Fermi screening for the screened-exchange energy;

$$E_{\text{sX}}[\{\psi\}] = -\frac{1}{2} \sum_{i,j}^{\text{occ}} \int \int d\mathbf{r} d\mathbf{r}' \frac{\psi_i^*(\mathbf{r}) \psi_j^*(\mathbf{r}') \psi_j(\mathbf{r}) \psi_i(\mathbf{r}') e^{-k_{\text{TF}}|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}, \quad (2)$$

where $\{\psi_i\}$ are the single particle orbitals that minimize the total energy in Eq.(1). We hereafter use atomic units, *i.e.* $\hbar = m = e = 1$, in equations through out our paper. The approximation in GKS arises in the evaluation of the unknown energy functional $R[\rho]$. Because the LDA is accurate for homogeneous electron systems, $R[\rho]$ is chosen so that the total energy of the homogeneous electron systems agrees with the LDA results;

$$R[\rho] = E_{\text{xc}}^{\text{LDA}}[\rho] - E_{\text{sX}}^{\text{loc}}[\rho]. \quad (3)$$

Here, $E_{\text{sX}}^{\text{loc}}[\rho]$ is the local density approximation of $E_{\text{sX}}[\{\psi\}]$;

$$E_{\text{sX}}^{\text{loc}}[\rho] = F[k_{\text{TF}}/k_F] E_{\text{x}}^{\text{LDA}}[\rho], \quad (4)$$

where $F[z]$ is a correction factor and k_F is the Fermi wave vector. It was suggested that the proper functional $E_{\text{sX}}^{\text{loc}}[\rho]$ should have the same density dependence as in $E_{\text{x}}^{\text{LDA}}$. [17] This is achieved by using a constant Fermi wave vector, k_F . To make the total energy agree with LDA in homogeneous systems, the average charge density, ρ_0 , is used for the Fermi and the Thomas Fermi wave vectors; $F_0 = F[k_{\text{TF}}/k_F]$ depends on only the average electron density, ρ_0 .

The self-consistent equations for the single particle orbitals are obtained by constraint-search minimization over single particle orbitals. The resulting GKS equations are

$$(T + V_{\text{ext}} + V_H + V_{\text{xc}}^{\text{LDA}}) \psi_{n\mathbf{k}}(\mathbf{r}) + \int d\mathbf{r}' V_{\text{sX}}(\mathbf{r}, \mathbf{r}') \psi_{n\mathbf{k}}(\mathbf{r}') - V_{\text{sX}}^{\text{loc}}(\mathbf{r}) \psi_{n\mathbf{k}}(\mathbf{r}) = \epsilon_{n\mathbf{k}} \psi_{n\mathbf{k}}(\mathbf{r}), \quad (5)$$

where T is the kinetic energy operator, V_{ext} is the external potential, V_H is the Hartree potential, and $V_{\text{xc}}^{\text{LDA}}$ is the LDA exchange-correlation potential. The first four terms on the left hand side of Eq. (5) are identical to the ones in the LDA KS equation. V_{sX} is the

non-local screened-exchange operator,

$$\int d\mathbf{r}' V_{\text{sX}}(\mathbf{r}, \mathbf{r}') \psi(\mathbf{r}') = - \sum_i^{\text{occ}} \psi_i(\mathbf{r}) \int d\mathbf{r}' \frac{\psi_i^*(\mathbf{r}') \psi(\mathbf{r}') e^{-k_{\text{TF}}|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}, \quad (6)$$

and the local potential $V_{\text{sX}}^{\text{loc}}$ is the functional derivative of $E_{\text{sX}}^{\text{loc}}$ with respect to the charge density $\rho(\mathbf{r})$,

$$V_{\text{sX}}^{\text{loc}}[\rho] = F_0 V_{\text{x}}^{\text{LDA}}[\rho]. \quad (7)$$

It is convenient to define an exchange charge density matrix,

$$\rho_{\text{x}}(\mathbf{r}, \mathbf{r}') = \sum_i^{\text{occ}} \psi_i(\mathbf{r}) \psi_i^*(\mathbf{r}') \quad (8)$$

so that the screened-exchange potential can be written as

$$\int d\mathbf{r}' V_{\text{sX}}(\mathbf{r}, \mathbf{r}') \psi(\mathbf{r}') = - \int d\mathbf{r}' \frac{\rho_{\text{x}}(\mathbf{r}, \mathbf{r}') e^{-k_{\text{TF}}|\mathbf{r}-\mathbf{r}'|} \psi(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \quad (9)$$

The advantage of sX-LDA is that the eigenvalue gaps of Eq.(5) agree better with experiments than the LDA eigenvalue gaps.[8, 17–20] Fig. 1 shows improved band gaps in sX-LDA for several tested zinc-blende structure semiconductors. The eigenfunctions of LDA and sX-LDA are, on the other hand, very similar. The overlap of corresponding LDA and sX-LDA eigenfunctions, $\langle \psi_{\mathbf{n}\mathbf{k}}^{\text{LDA}} | \psi_{\mathbf{n}\mathbf{k}}^{\text{sX-LDA}} \rangle$, is bigger than 99.9% for all bands and Brillouin zone k -points, and different state eigenfunctions are almost mutually orthonormal, *i.e.*, $\langle \psi_{\mathbf{n}\mathbf{k}}^{\text{LDA}} | \psi_{\mathbf{n}'\mathbf{k}'}^{\text{sX-LDA}} \rangle \approx \delta_{n,n'} \delta_{\mathbf{k},\mathbf{k}'}$, in all of our testing cases. This observation is analogous to the well known fact that single particle wavefunctions calculated from Hartree-Fock and GW method are very similar to those from LDA method.[21]

The calculations that are presented in the paper were done by using a plane-wave sX-LDA code implemented in PEtot code[22] with Troullier-Martins norm-conserving pseudopotentials.[23] For all investigated semiconductors, we used 19 special \mathbf{k} points in the irreducible wedge of Brillouin zone[24] to evaluate the \mathbf{k} summation in the electron density and the summation of the non-local potential in Eq.(6). The plane-wave basis was cut off at a kinetic energy of 50 Ry for diamond, 30 Ry for Si and GaAs, and 20 Ry for Ge. For LDA correlation energy, we used Perdew-Zunger parametrization of the electron gas correlation energy.[25]

III. SINGLE-PARTICLE POTENTIALS: COMPARISON WITH LDA AND GW

In this section we investigate sX-LDA formalism by comparing the eigenvalues of GKS equations and the exchange-correlation potential of sX-LDA with the quasi-particle energies and the self-energy of the GW method. We look into differences in eigenvalue gaps as well as absolute eigenvalues of LDA, sX-LDA, and GW methods. We also study the nonlocality of sX-LDA and GW potentials.

The quasi-particle GW equation in many-body perturbation theory is

$$(T + V_{\text{ext}} + V_H)\psi_{n\mathbf{k}}(\mathbf{r}) + \int d\mathbf{r}' \Sigma(\mathbf{r}, \mathbf{r}'; \epsilon_{n\mathbf{k}})\psi_{n\mathbf{k}}(\mathbf{r}') = \epsilon_{n\mathbf{k}}\psi_{n\mathbf{k}}(\mathbf{r}) . \quad (10)$$

The difference in the GW quasi-particle equation from the GKS equation is the energy dependent nonlocal self-energy term, $\Sigma(\mathbf{r}, \mathbf{r}'; \epsilon_{n\mathbf{k}})$. Similarly to the sX-LDA eigenfunctions, the eigenfunctions of GW quasiparticle equations are very close to the eigenfunctions of LDA KS equations.[21] Based on this similarity in eigenfunctions, we focus on the exchange-correlation potential in the single-particle equations. Given the same charge density, the other terms in single particle equations, *i.e.*, the kinetic, Hartree, and external potentials, are identical in LDA, sX-LDA and GW. Specifically, we analyze the exchange-correlation potential matrix elements of (sX-)LDA and the self-energy matrix elements of GW.

We first show the fundamental energy gap of the test systems in LDA, sX-LDA, and GW method in Table I. It has been reported that LDA and sX-LDA yield different lattice constants and that sX-LDA improves the lattice constant closer to experiments in most cases.[8] In all calculations reported here, both LDA and sX-LDA, we used experimental lattice constants.[26] In general, the calculated sX-LDA band gap is in good agreement with experiments. The agreement of sX-LDA and GW with experiments varies case by case. We have not found definite relationship between the energy band gaps in sX-LDA and GW. Compared with GW predictions, sX-LDA band gap is better for diamond but worse for silicon while the band gap predictions are identical for germanium. Note that sX-LDA tends to overestimate the X point conduction band energy.

Because the eigenfunctions of LDA and sX-LDA are almost identical, the correction in the eigenvalue gap can be directly attributed to the exchange-correlation potential matrix element, $\langle n\mathbf{k} | V_{\text{xc}} | n\mathbf{k} \rangle$. For LDA, V_{xc} is the usual local potential $V_{\text{xc}}^{\text{LDA}}$. For sX-LDA, V_{xc} is the sum of the local $V_{\text{xc}}^{\text{LDA}}$, the nonlocal screened-exchange operator, V_{sX} , and minus its

local approximation, V_{sX}^{loc} ;

$$V_{xc}^{\text{sX-LDA}}(\mathbf{r}, \mathbf{r}') = \{(1 - F_0)V_x^{\text{LDA}}(\mathbf{r}) + V_c^{\text{LDA}}(\mathbf{r})\}\delta(\mathbf{r} - \mathbf{r}') - \frac{\rho_x(\mathbf{r}, \mathbf{r}')e^{-k_{TF}|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|}. \quad (11)$$

The effects of the screened-exchange potential on the eigenvalues can be seen by evaluating the exchange-correlation matrix elements in LDA eigenfunction basis. Within a perturbation theory, the first order sX-LDA correction in eigenvalues is the expectation value of $\Delta V_{xc}(\mathbf{r}, \mathbf{r}') = V_{xc}^{\text{sX-LDA}}(\mathbf{r}, \mathbf{r}') - V_{xc}^{\text{LDA}}(\mathbf{r}, \mathbf{r}')$ (difference of exchange-correlation potential between LDA and sX-LDA);

$$\epsilon_{n,\mathbf{k}}^{\text{sX-LDA}} \approx \epsilon_{n,\mathbf{k}}^{\text{LDA}} + \langle \psi_{n,\mathbf{k}}^{\text{LDA}} | \Delta V_{xc} | \psi_{n,\mathbf{k}}^{\text{LDA}} \rangle. \quad (12)$$

In Table II, we compare the self-consistent band gap results with the first order estimation of the band gaps using LDA wavefunctions. The good agreement between self-consistent band gaps and perturbative estimations of the band gap has been reported in previous studies.[8] Small disagreement arises from the charge density change in the self-consistent results, which usually shows an increase of charge density in the bonding region in sX-LDA.

To clarify the origin of the improvement in band gaps, we show the difference between the self-consistent eigenvalues obtained from sX-LDA and LDA as a function of the eigenvalues of sX-LDA in Fig. 2. In all the cases we studied, there is a big jump in the eigenvalue difference from valence band maximum (VBM) to conduction band minimum (CBM). Away from the band gap, the difference increases monotonically with the energy difference from the VBM energy, which results in the stretch of bands rather than a rigid shift of bands. These results are similar to those of the GW calculations.[21]

Another feature in Fig.(2) is the downshift of valence band eigenvalues. Within DFT formalism, the VBM eigenvalue is the minus ionization energy.[27] The increases in the ionization energy in sX-LDA calculation is 1.84, 1.16, 0.86, and 0.97 eV for diamond, Si, Ge, and GaAs, respectively. This increased ionization energy is due to reduced self-interactions in sX-LDA. A similar increase in ionization energy is observed in self-interaction corrected DFT (SIC-DFT).[25] Note that, unlike SIC-DFT where the self-interaction is completely removed, in sX-LDA the self-interaction is removed only at short-range where $r \lesssim k_{TF}^{-1}$.

In Fig. 3 we plot the exchange-correlation potential matrix elements of VBM and CBM for diamond, Si, and GaAs. The matrix elements are evaluated using LDA wavefunctions. Going from LDA to sX-LDA, the matrix elements are down-shifted for both CBM and VBM.

The magnitude of the changes in VBM is much bigger than in CBM. Therefore, the band gap correction in sX-LDA primarily comes from the deepening of VBM state energy. This is a marked difference from the GW results, where the self-energy matrix elements of GW are very similar to LDA exchange-correlation potential matrix elements for VBM but up-shifted for CBM. On the other hand, our finding in sX-LDA is rather similar to the results of Coulomb-hole-screened-exchange (COHSEX) approximation of the GW self-energy term. [21]

To probe the difference between sX-LDA and GW more directly, we look into the non-locality of the exchange-correlation potentials. In case of the GW method, the self-energy term depends on the quasi-particle energy. However, we found that the self-energy term has a weak dependence near the band gap and remains almost unchanged when the quasi-particle energy changes from VBM to CBM. For the convenience of discussion, we fix the quati-particle energy to VBM energy. The three dimensional nonlocal potential $V_{xc}^{sX-LDA}(\mathbf{r}, \mathbf{r}')$ is difficult to visualize in a quantitative way. We plot the spherical average of the exchange-correlation potential multiplied by the square of the radial distance, which provides a proper weight for radial energy density. More specifically, we use

$$V_{xc}(\mathbf{r}, R)R^2 = R^2 \int \frac{d\Omega_R}{4\pi} V_{xc}(\mathbf{r}, \mathbf{r} + \mathbf{R}), \quad (13)$$

where one electron is fixed at a position \mathbf{r} . The nonlocality can be seen from this spherically average potential and the anisotropy of the original potential. In sX-LDA, the exponentially decaying nonlocal potential hardly shows any anisotropy. Fig.4 shows $V_{xc}(\mathbf{r}, R)R^2$ of Si with \mathbf{r} at bond center, atom, and interstitial sites, along with GW and HF nonlocal potentials. Note that the local part of $V_{xc}^{sX-LDA}(\mathbf{r}, \mathbf{r}')$ in Eq(11) gives rise to a delta peak in $V_{xc}(\mathbf{r}, R)$ at $R = 0$. This delta peak is not shown in Fig.4, when multiplied by R^2 , but should be taken into account when we estimate the overall amplitude and the nonlocal range of the the exchange-correlation potential, especially when we compare the sX-LDA exchange-correlation potential to the GW and HF nonlocal potentials, which do not have such delta peak. The local part of $V_{xc}(\mathbf{r}, \mathbf{r}')$ are -2.00, -1.51, and -1.07 eV at bond center, atom, and interstitial site, respectively.

Among the sX-LDA, GW and HF nonlocal potentials, HF has the largest and longest range nonlocal potential, followed by the GW nonlocal potential. Compared with HF non-local potential, GW potential is screened in the long range side, and its center of mass is

shifted towards the short range side. The amplitude of sX-LDA nonlocal potential is too small, especially for the small density region, e.g., the interstitial site. In addition, when the delta peak at the origin is taken into account, the range of nonlocality is also too small compared with the GW potentials. This implies that the range of the screening should be increased to increase the nonlocality of the sX-LDA. The relative difference in the magnitude of the potential is more pronounced at the low electron density region, e.g., atom and the interstitial sites. The dependence of sX-LDA and GW discrepancy on the local charge density indicates that the homogeneous screening function with a constant screening length in sX-LDA is inadequate to describe the nonlocality of the exchange-correlation potentials. It might also contribute to the fact that sX-LDA often has a too high conduction band state energy at zinc-blende X point, where the wavefunction has large amplitude at the interstitial region.

IV. EXCHANGE-CORRELATION HOLE: COMPARISON WITH LDA AND QMC

In this section we investigate the effects of the nonlocal screened-exchange potential on total energy calculations. Unlike the GW method, where the interest is solely in the quasi-particle energies, the screened-exchange method also provides the total energy, thus yielding information on the ground state configuration. While sX-LDA method is constructed to have the same total energy as in the LDA method for homogeneous systems, the comparison between LDA and sX-LDA total energy via the equilibrium lattice constants and bulk moduli of semiconductors indicates that the two methods can be substantially different in real systems. The sX-LDA results are better for the lattice constants but worse in bulk moduli. [8] Both of those quantities, however, give information on the relative total energy for different lattices and atomic structures, but they do not serve as the measure of the accuracy of the absolute total energy. Because of the similarity of wavefunctions in LDA and sX-LDA, the majority of the total energy difference in the two method arises from the exchange-correlation energy. Since the comparison of the total energy alone might not provide enough information on the the origin of their differences, it is more interesting to study the exchange-correlation hole itself.

To properly define the exchange-correlation hole in different formalisms, we start with

the exact exchange-correlation energy expression.

$$E_{\text{xc}}[n(\mathbf{r})] = \frac{1}{2} \int d\mathbf{r} \rho(\mathbf{r}) \int d\mathbf{r}' \frac{n_{\text{xc}}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \quad (14)$$

where the exchange-correlation hole n_{xc} can be related to the symmetric pair distribution function, $g(\mathbf{r}, \mathbf{r}')$;

$$n_{\text{xc}}(\mathbf{r}, \mathbf{r}') = \rho(\mathbf{r}') [g(\mathbf{r}, \mathbf{r}') - 1]. \quad (15)$$

Similarly, the exchange-correlation hole, $n_{\text{sX}}(\mathbf{r}, \mathbf{r}')$, due to the nonlocal part of screened-exchange term in the sX-LDA formalism is

$$E_{\text{sX}} = \frac{1}{2} \int d\mathbf{r} \rho(\mathbf{r}) \int d\mathbf{r}' \frac{n_{\text{sX}}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \quad (16)$$

where

$$n_{\text{sX}}(\mathbf{r}, \mathbf{r}') = -\frac{|\rho_{\text{x}}(\mathbf{r}, \mathbf{r}')|^2 e^{-k_{\text{TF}}|\mathbf{r}-\mathbf{r}'|}}{\rho(\mathbf{r})}. \quad (17)$$

Here $\rho_{\text{x}}(\mathbf{r}, \mathbf{r}')$ is defined in Eq.(8). Note that the corresponding contribution to the pair distribution function

$$g_{\text{sX}}(\mathbf{r}, \mathbf{r}') = 1 + \frac{n_{\text{sX}}(\mathbf{r}, \mathbf{r}')}{\rho(\mathbf{r}')} \quad (18)$$

is symmetric with particle exchange between \mathbf{r} and \mathbf{r}' . The full sX-LDA exchange-correlation hole has also contributions from the local density functional parts. The total sX-LDA exchange-correlation hole is

$$n_{\text{xc}}^{\text{sX-LDA}}(\mathbf{r}, \mathbf{r}') = (1 - F_0) n_{\text{x}}^{\text{LDA}}(\mathbf{r}, \mathbf{r}') + n_{\text{c}}^{\text{LDA}}(\mathbf{r}, \mathbf{r}') + n_{\text{sX}}(\mathbf{r}, \mathbf{r}'). \quad (19)$$

Here, following the convention, the LDA exchange and correlation hole, $n_{\text{x}}^{\text{LDA}}(\mathbf{r}, \mathbf{r}')$ and $n_{\text{c}}^{\text{LDA}}(\mathbf{r}, \mathbf{r}')$, can be defined as

$$\begin{aligned} n_{\text{x}}^{\text{LDA}}(\mathbf{r}, \mathbf{r}') &= U_{\text{x}}(|\mathbf{r} - \mathbf{r}'| \rho(\mathbf{r})^{1/3}), \\ n_{\text{c}}^{\text{LDA}}(\mathbf{r}, \mathbf{r}') &= U_{\text{c}}(|\mathbf{r} - \mathbf{r}'| \rho(\mathbf{r})^{1/3}), \end{aligned} \quad (20)$$

where $U_{\text{x}}(x)$ and $U_{\text{c}}(x)$ are universal functions derived from the exchange and correlation energy expressions of homogeneous electron systems.[25] We point out that, in contrast to $n_{\text{sX}}(\mathbf{r}, \mathbf{r}')$, the corresponding pair-distribution functions of $n_{\text{x}}^{\text{LDA}}(\mathbf{r}, \mathbf{r}')$ and $n_{\text{c}}^{\text{LDA}}(\mathbf{r}, \mathbf{r}')$ are not symmetric in particle exchange. As a result, the LDA exchange-correlation hole does not connect the physical properties of the two points, \mathbf{r} and \mathbf{r}' . In this sense it is not the true nonlocal hole. LDA XC hole is only defined by local properties (charge density) at

\mathbf{r} . Nevertheless, it is useful to compare $n_{\text{xc}}^{\text{LDA}}(\mathbf{r}, \mathbf{r}')$ with the exchange correlation hole from other methods since the LDA exchange-correlation energy can be calculated from $n_{\text{xc}}^{\text{LDA}}(\mathbf{r}, \mathbf{r}')$.

Unlike in the Hartree-Fock approximation, the exchange-correlation sum rule is not fully satisfied in the screened-exchange method, *i.e.*, $\int d\mathbf{r}' n_{\text{xc}}^{\text{sX-LDA}}(\mathbf{r}, \mathbf{r}') \neq -1$. For example, the total exchange-correlation hole in Si with one electron fixed at bond center, atom site, and interstitial site are -0.996, -0.928, and -0.859, respectively. Like the nonlocal potential discussed in previous section, the three dimensional nonlocality of the exchange-correlation hole is hard to visualize, and it is often easier to see the characteristics of the functions when the spherical average of the functions is used. For this regards, we define the spherically averaged exchange-correlation hole as

$$\tilde{n}_{\text{xc}}(\mathbf{r}, R) = \int \frac{d\Omega_R}{4\pi} n_{\text{xc}}(\mathbf{r}, \mathbf{r} + \mathbf{R}) \quad (21)$$

and similarly, $\tilde{n}_{\text{x}}(\mathbf{r}, R)$ and $\tilde{n}_{\text{sX}}(\mathbf{r}, R)$ for spherically averaged exact exchange and screened-exchange hole, respectively.

We now look into the effects of the screening on the exchange hole. In Fig. 5 we show the screened and unscreened Hartree-Fock exchange holes with one electron fixed at the bond center, the atom site, and the interstitial site using LDA wavefunctions. The exact Hartree-Fock exchange hole shows more structure reflecting the electron density variation in the real space. On the contrary, the screening in sX-LDA suppresses the long-range nonlocal feature of the exchange hole, and n_{sX} contributes only in short-range. Note the difference in the magnitude of the exchange hole in different position. The absolute difference is larger in bonding region while the relative difference is larger in lower density region.

The exchange-correlation hole in the sX-LDA method deviates from the LDA counterpart due to the short-range nonlocal screened-exchange potential. In Fig. 6, we show the spherically averaged exchange-correlation hole, $\tilde{n}_{\text{xc}}(\mathbf{r}, R)$, from the sX-LDA method and LDA method with one electron fixed at the same positions as in Fig. 5. For LDA calculations, we used the self-consistent sX-LDA charge density and the Perdew-Wang pair-distribution function. [28] We also show the exchange-correlation hole from the coupling constant-averaged variational Monte Carlo (VMC) calculations in Ref. [16] for qualitative comparison. Compared with LDA method, the sX-LDA exchange-correlation hole is larger at short distance while smaller at long distance, regardless of the position of the origin \mathbf{r} . This can also be seen from the formula. Since there is no screening at $\mathbf{r} = \mathbf{r}'$, the exchange-correlation hole

is $(2 - F_0)n_x^{\text{LDA}}(\mathbf{r}, \mathbf{r}) + n_c^{\text{LDA}}(\mathbf{r}, \mathbf{r})$, thus increasing the exchange hole by $(1 - F_0)n_x^{\text{LDA}}(\mathbf{r}, \mathbf{r})$ relative to LDA. Comparison with VMC results reveals that the LDA gives better exchange-correlation hole than sX-LDA, particularly in high electron density region. The Hartree-Fock exchange holes in Fig. 5 (which are also replotted in Fig. 6 show similar structures as in VMC exchange-correlation hole. Similar as in the case of the nonlocal potential in Fig. 4 (c) at the small density region, when compared to VMC, the HF exchange-correlation hole structure has been enhanced in the short distance range, while compressed in the long distance range. Overall, the comparison indicates that the Thomas-Fermi screening in the current sX-LDA method is too short range, and a constant screening length for all points in the system is inadequate. These conclusions are similar to the conclusions of the previous section based on the potential consideration.

Finally, we relate the exchange-correlation hole to the total energy calculation. Using the spherically averaged exchange-correlation hole, the total exchange-correlation energy density at \mathbf{r} can be expressed as

$$e_{\text{xc}}(\mathbf{r}) = \frac{1}{2} \int d\mathbf{r}' \frac{n_{\text{xc}}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = 2\pi \int dR R \tilde{n}_{\text{xc}}(\mathbf{r}, R). \quad (22)$$

Because the above integration has a factor R , the contribution of \tilde{n}_{xc} from the long range weighs more in the energy calculation. In Fig. 7 we show the distance weighted exchange-correlation hole $2\pi\tilde{n}_{\text{xc}}(\mathbf{r}, R)$. Although the structure of n_x and n_{sX} is more different in lower density region, the contribution to the total energy difference is larger at higher density region. Table IV shows $e_{\text{xc}}(\mathbf{r})$ at the bond center and the interstitial site. Compared with other methods, sX-LDA has the largest disagreement with VMC. The absolute disagreement in the exchange-correlation energy density, $\rho(\mathbf{r})e_{\text{xc}}(\mathbf{r})$, is larger at the bond center.

In Table III, we compare the total exchange-correlation energy calculated from $\int e_{\text{xc}}(\mathbf{x})\rho(\mathbf{r})d\mathbf{r}$ in LDA and sX-LDA, and the QMC results from the Ref[16]. The small energy difference ≈ 0.15 eV in the two LDA results can be attributed to the different pseudopotentials. Clearly, the discrepancy of the sX-LDA exchange-correlation energy from diffusion Monte Carlo (DMC) result is larger than that of LDA because of the short-range overestimation of the exchange hole. The large difference between sX-LDA and DMC, ≈ 1.0 eV, is beyond the pseudopotential error. Thus, in terms of the absolute total energy, the LDA method is more accurate than the sX-LDA method.

V. CONCLUSION

To summarize, we have the following conclusions: (1) The improvement of eigenvalue gap in sX-LDA has different origin from the GW method. In sX-LDA the gap increases is due to the deepening of the VBM state energy, while in GW method it is primarily due to the increase of the CBM state energy. (2) In comparison with the QMC result, the absolute sX-LDA exchange-correlation energy is too negative, and it is worse than the LDA results. (3) The sX-LDA exchange-correlation nonlocal potential is much smaller than the GW self-energy term, and the sX-LDA potential has a big local component. Thus, overall, the nonlocality of the sX-LDA nonlocal potential is not large enough, and it is too short range. (4) In comparison with the QMC result, the sX-LDA exchange-correlation hole is overestimated in the short range, mainly due to the local density approximation part of the exchange-correlation hole. Further more, it lacks the oscillating structure at the small charge density sites (atom and interstitial sites) which exists in both the HF and QMC XC holes. (5) The uniform screening for all sites with the constant screening length in the current sX-LDA formalism might not be appropriate. This might contribute to the fact that the sX-LDA tends to overestimate the conduction band energy at the X point.

VI. ACKNOWLEDGEMENTS

BL acknowledges R. M. Martin and W. Ku for fruitful discussions. BL and LWW were supported by U.S. Department of Energy under contract No. DE-AC03-76SF00098. CDS was partially supported by the Nanoscale Science and Engineering Initiative of the National Science Foundation under NSF Award Numbers CHE-0117752 and CHE-0641523, and by the New York State Office of Science, Technology, and Academic Research (NYSTAR). The calculations were done using resources of NERSC at Lawrence Berkeley National Lab.

-
- [1] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
 - [2] J. P. Perdew and Y. Wang, Phys. Rev. B **33**, 8800 (1986).
 - [3] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
 - [4] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).
 - [5] L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).
 - [6] J. F. Janak, Phys. Rev. B **18**, 7165 (1978).
 - [7] L. J. Sham and M. Schlüter, Phys. Rev. B **32**, 3883 (1985).
 - [8] A. Seidl, A. Gorling, P. Vogl, J. A. Majewski, and M. Levy, Phys. Rev. B **53**, 3764 (1996).
 - [9] L. Hedin, Phys. Rev. **139**, A796 (1965).
 - [10] L. Hedin and S. Lundqvist, *Effects of Electron-Electron and Electron-Phonon Interactions on the One-Electron States of Solids*, vol. 23 of *Solid State Physics* (Academic, New York, 1969).
 - [11] G. Strinati, H. J. Mattausch, and W. Hanke, Phys. Rev. Lett. **45**, 290 (1980).
 - [12] M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985).
 - [13] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. Lett. **56**, 2415 (1986).
 - [14] F. Aryasetiawan and O. Gunnarsson, Rep. Prog. Phys. **61**, 237 (1998).
 - [15] W. Aulbur, L. Jönsson, and J. Wilkins, *Quasiparticle Calculations in Solids*, vol. 54 of *Solid State Physics* (Academic, New York, 2000).
 - [16] R. Q. Hood, M. Y. Chou, A. J. Williamson, G. Rajagopal, and R. J. Needs, Phys. Rev. B **57**, 8972 (1998).
 - [17] D. M. Bylander and L. Kleinman, Phys. Rev. B **41**, 7868 (1990).
 - [18] C. B. Geller, W. Wolf, S. Picozzi, A. Continenza, R. Asahi, W. Mannstadt, A. J. Freeman, and E. Wimmer, Appl. Phys. Lett. **79**, 368 (2001).
 - [19] B. Lee and L.-W. Wang, Phys. Rev. B **73**, 153309 (2006).
 - [20] B. Lee and L.-W. Wang, J. Appl. Phys. **100**, 093717 (2006).
 - [21] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
 - [22] URL <http://hpcrd.lbl.gov/~linwang/PEtot/PEtot.html>.
 - [23] N. Troullier and J. Martins, Solid State Commun. **74**, 613 (1990).
 - [24] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).
 - [25] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

- [26] O. Madelung, M. Schulz, and H. Weiss, eds., *Intrinsic Properties of Group IV Elements and III-V, II-VI, and I-VII Compounds*, vol. 22 Pt. a of *Landolt-Bornstein, New Series* (Springer, Berlin, 1987).
- [27] J. C. Slater, Adv. Quantum Chem. **6**, 1 (1972).
- [28] J. P. Perdew and Y. Wang, Phys. Rev. B **46**, 12947 (1992).

TABLE I: The fundamental band gaps (in eV) of diamond, silicon, germanium, and GaAs calculated from LDA, sX-LDA, and GW are compared to experiments.

	LDA	sX-LDA	GW	Experiment
Diamond	4.20	5.41	5.6 ^a	5.41 ^b
Silicon	0.52	1.42	1.29 ^a	1.17 ^b
Germanium	< 0	0.74	0.75 ^a	0.74 ^b
GaAs	0.62	1.46		1.52 ^b

^aReference 21

^bReference 26

TABLE II: Comparison of the fundamental band gaps difference and the exchange-correlation potential matrix element difference. $\Delta E_g = E_g^{\text{sX-LDA}} - E_g^{\text{LDA}}$; the self-consistent band gap difference. $\Delta E_g^{(1)} = \langle \psi_{\text{CBM}}^{\text{LDA}} | \Delta V_{\text{xc}} | \psi_{\text{CBM}}^{\text{LDA}} \rangle - \langle \psi_{\text{VBM}}^{\text{LDA}} | \Delta V_{\text{xc}} | \psi_{\text{VBM}}^{\text{LDA}} \rangle$; the first order perturbation estimation of the band gap. All energies in eV.

	Diamond	Silicon	GaAs
ΔE_g	1.21	0.90	0.84
$\Delta E_g^{(1)}$	1.03	0.96	0.89

TABLE III: Exchange-correlation energy (in eV) of bulk Si. The QMC results (VMC and DMC) are coupling strength averaged.

sX-LDA	LDA	LDA ^a	VMC ^a	DMC ^a
-34.19	-32.59	-32.75	-32.73 \pm 0.01	-33.23 \pm 0.08

^aReference 16

TABLE IV: Exchange-correlation energy density per particle (in atomic unit) of bulk Si. For LDA and HF results, the sX-LDA wavefunctions were used. The numbers inside the parentheses are the relative difference (in %) from the VMC results.

	bond center	interstitial
sX-LDA	-0.425 (-14.6)	-0.119 (28.3)
HF	-0.345 (7.0)	-0.152 (8.4)
LDA	-0.391 (-5.4)	-0.130 (21.7)
LDA ^a	-0.381 (-2.7)	-0.133 (19.9)
VMC ^a	-0.371	-0.166

^aReference 16

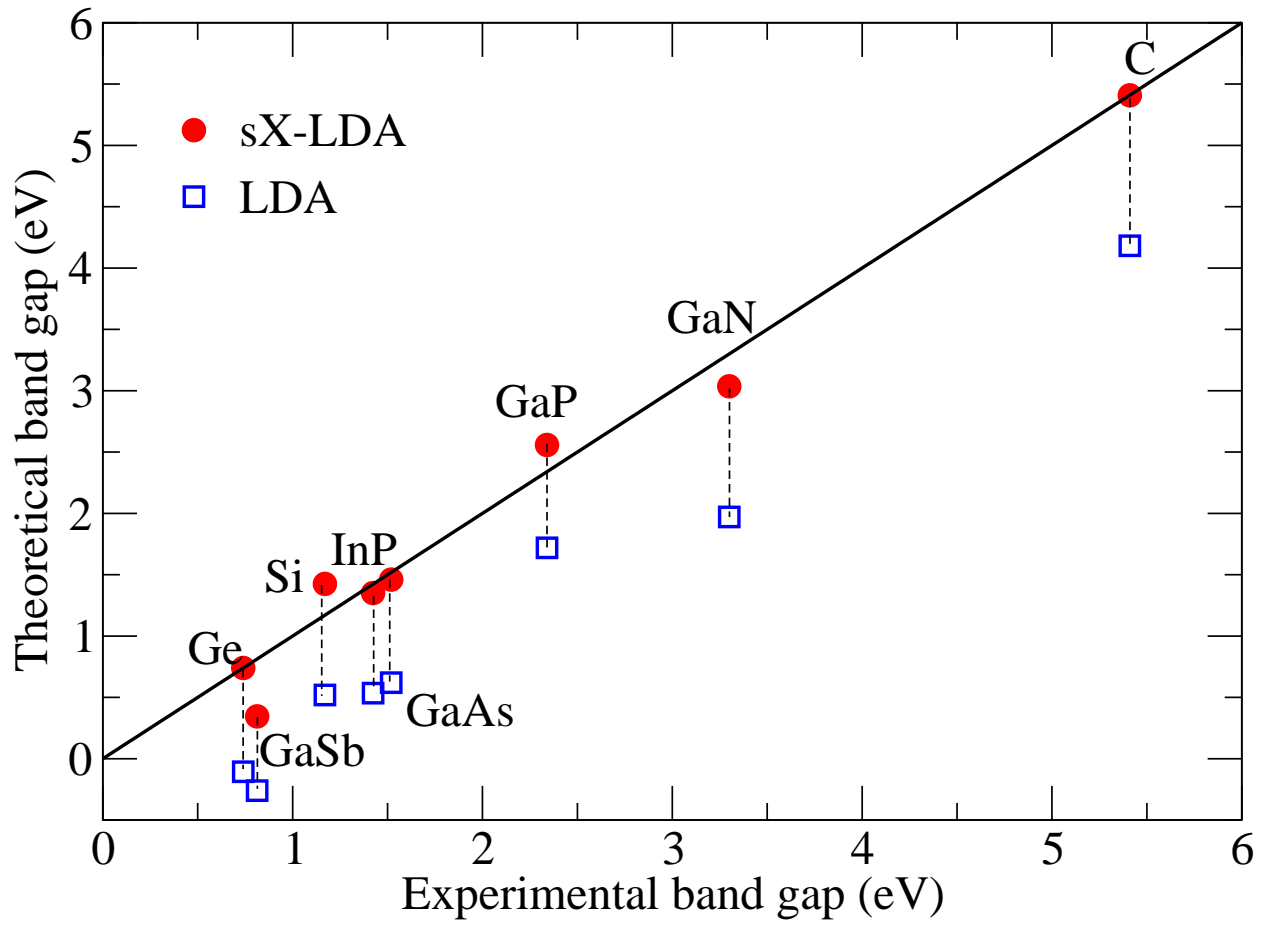


FIG. 1: (Color online) LDA (blue open squares) and sX-LDA (red filled circles) band gap *v.s.* the experimental value for several semiconductors in the zinc-blende structure.

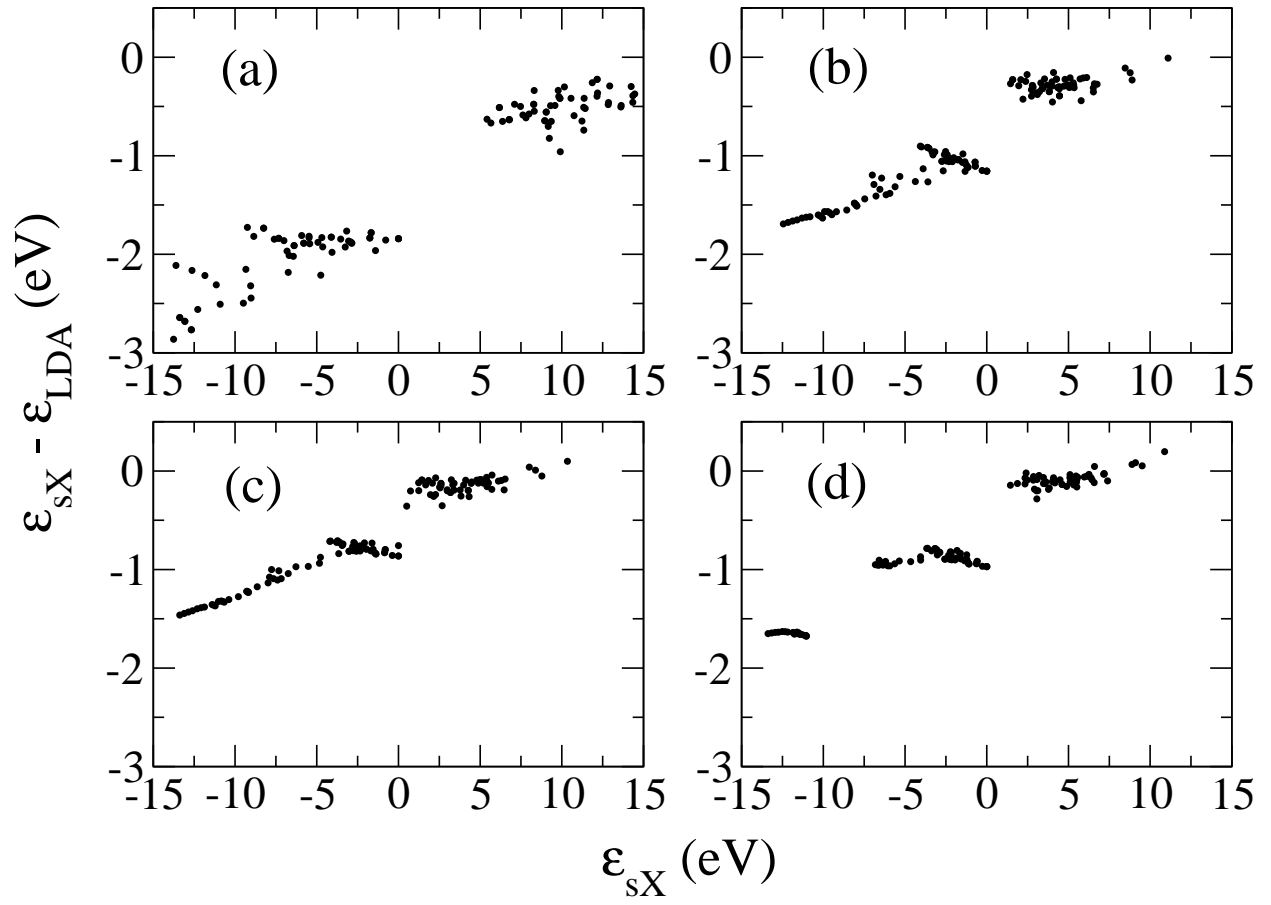


FIG. 2: The eigenvalue difference between sX-LDA and LDA, $\epsilon_{nk}^{LDA} - \epsilon_{nk}^{sX-LDA}$, is plotted against the sX-LDA eigenvalue for several symmetric Brillouin zone k -points: (a) diamond, (b) Si, (c) Ge, and (d) GaAs. The zero of the eigenvalue is set at the top of the valence band maximum.

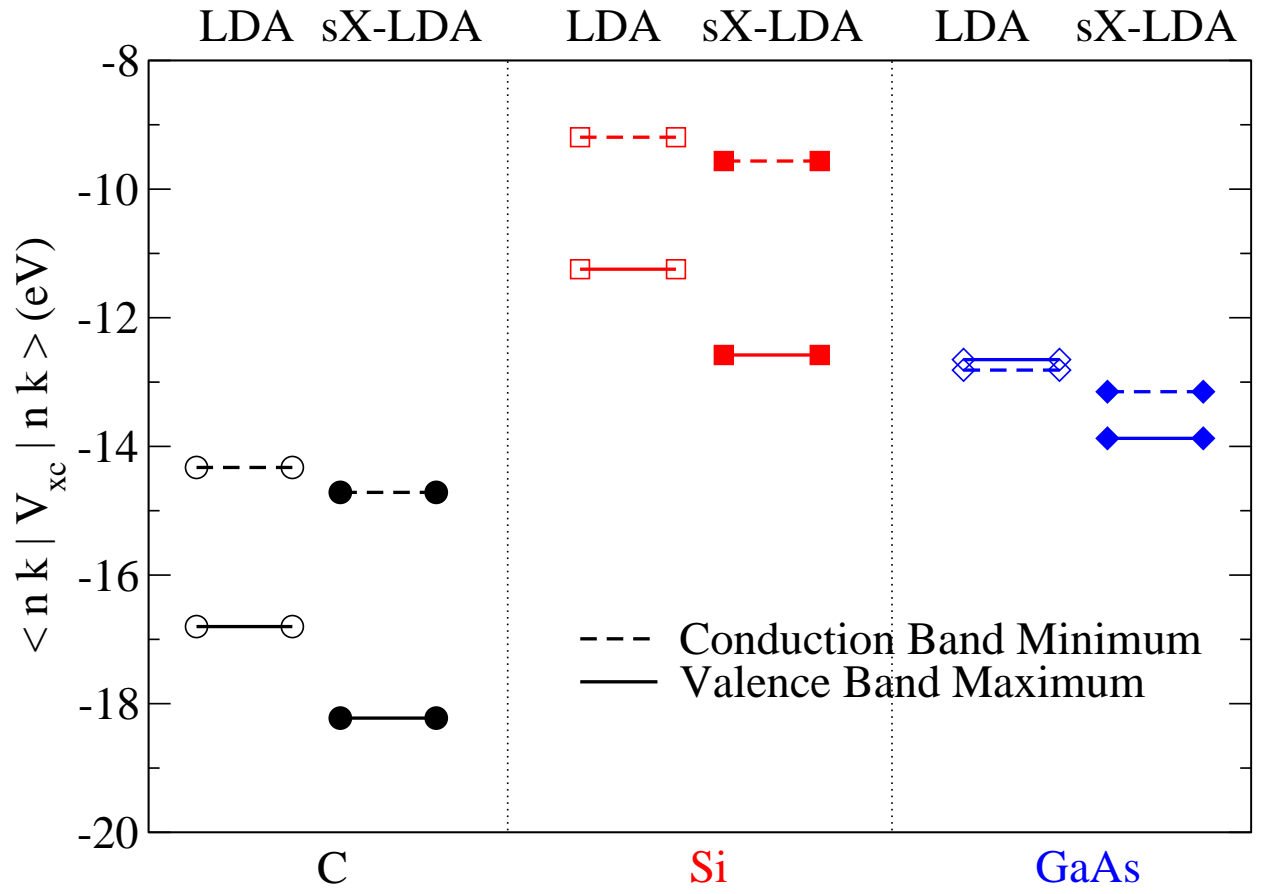


FIG. 3: (Color online) Matrix elements of exchange-correlation potential, V_{xc} , evaluated using the LDA wavefunctions. The matrix elements of the conduction band minimum and the valence band maximum are shown.

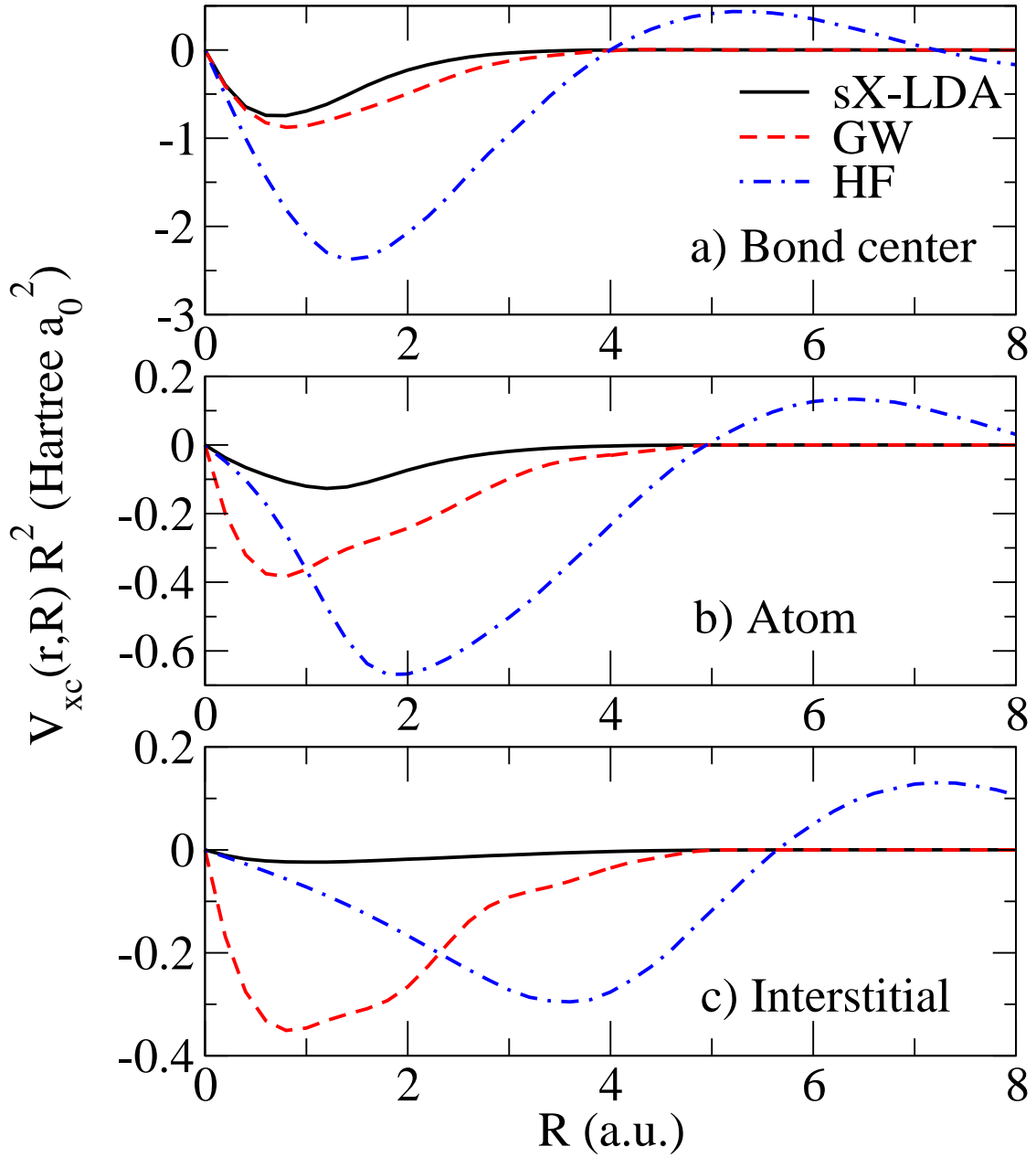


FIG. 4: (Color online) $V_{xc}(\mathbf{r}, R)R^2$. One electron is fixed at a) bond center, b) atom, and c) the interstitial site. The solid (black), dashed (red), dashed-dotted (blue) lines correspond to sX-LDA, GW, and HF results. The raw GW data show oscillation owing to the finite wavevector cutoff used in our calculations. The data in the current figure were obtained by taking the median values of the oscillating curves. For HF results, we used sX-LDA wavefunctions. Note that the figure shows only the *non-local* part. For sX-LDA, there is a sharp peak in $V_{xc}(\mathbf{r}, R)$ at $R = 0$ due to the local part of the exchange-correlation potentials. This delta peak is not shown in the figure when $V_{xc}(\mathbf{r}, R)$ is multiplied by R^2 .

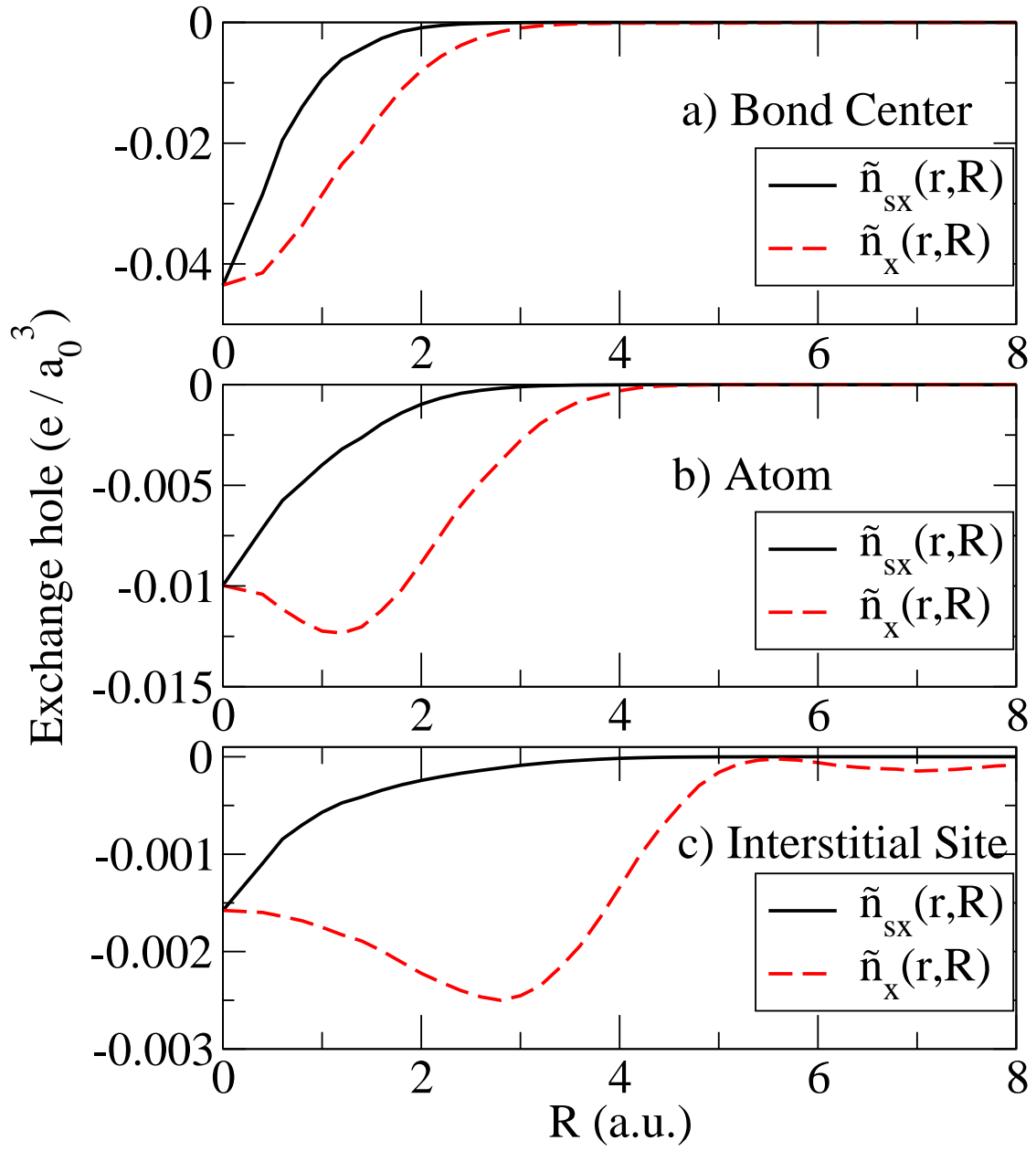


FIG. 5: (Color online) The spherically averaged screened (black solid line) and Hartree-Fock (red dashed line) exchange hole of bulk silicon are shown with one electron fixed at three different spatial points: a) bond center, b) atom, and c) interstitial site. The exchange holes were calculated using LDA wavefunctions.

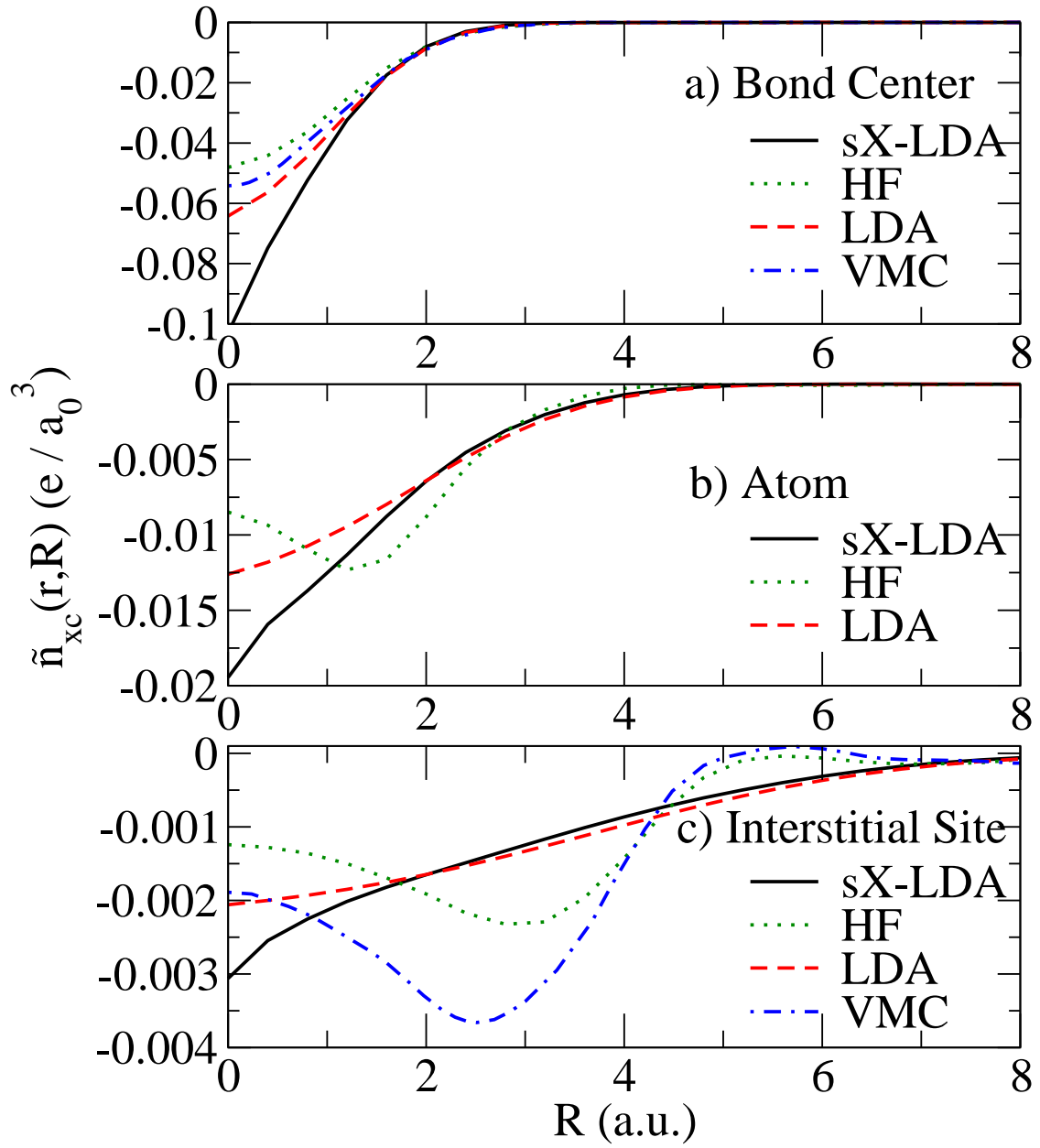


FIG. 6: (Color online) The spherically averaged sX-LDA exchange-correlation hole of bulk silicon in the screened-exchange method (black solid lines) is shown with one electron is fixed at three different spatial points: a) bond center, b) atom, and c) interstitial site. For comparison purpose, we also show the HF (green dotted lines), LDA (red dashed lines), and coupling strength-averaged VMC (blue dashed-dotted lines) exchange-correlation hole. The LDA and HF exchange-correlation holes were calculated using sX-LDA wavefunctions. The VMC results are from Ref. [16].

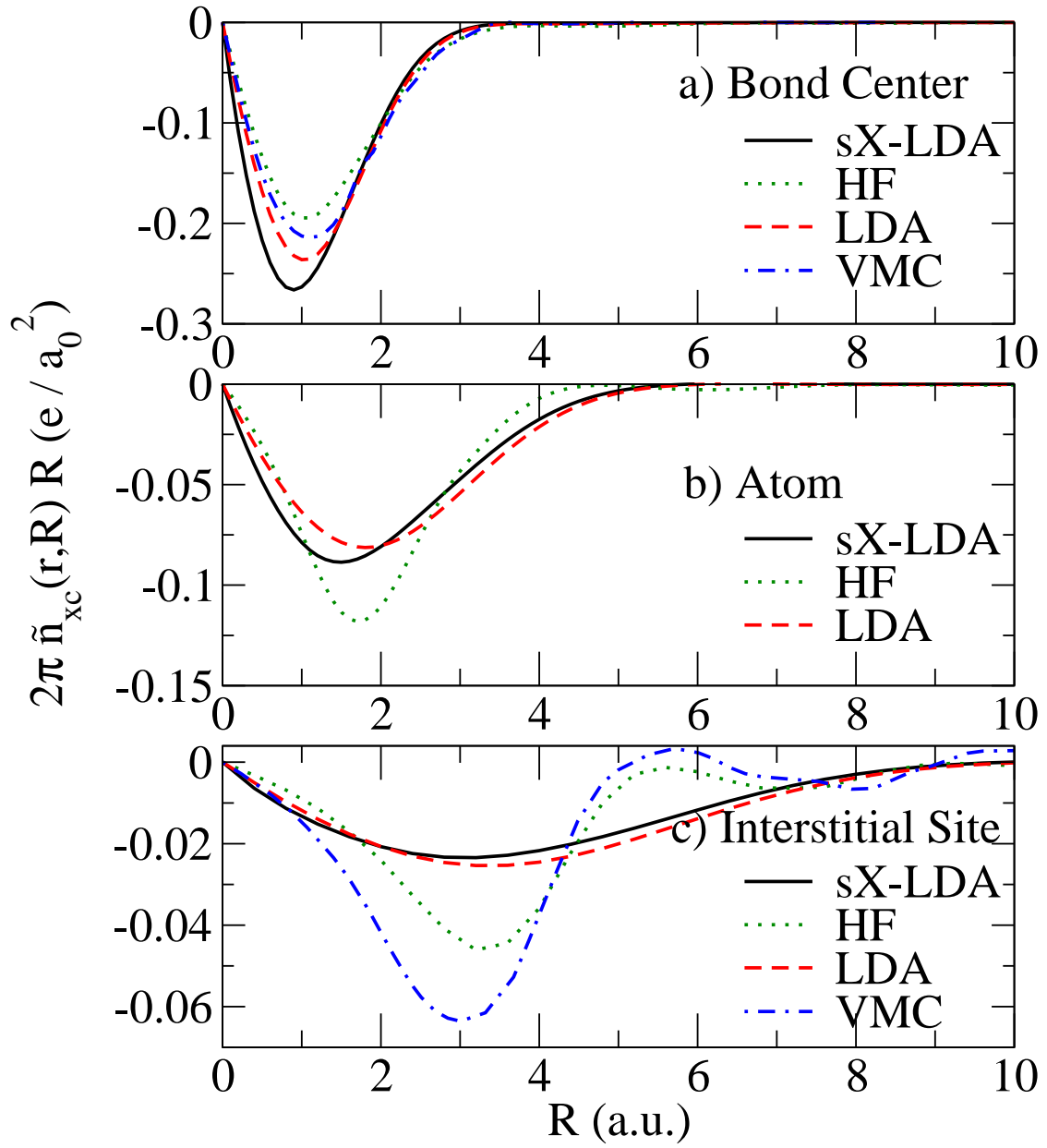


FIG. 7: (Color online) The distance weighted exchange-correlation hole of bulk silicon in the screened-exchange method (black solid lines) is shown with one electron is fixed at three different spatial points: a) bond center, b) atom, and c) interstitial site. The LDA (red dashed lines) and HF (green dotted lines) exchange-correlation holes were calculated using sX-LDA wavefunctions for comparison purpose. The VMC (blue dashed-dotted lines) results are from Ref. [16].